

**Interactive computer system e.g. mouse with hand gesture controlled operation - has 2 or 3 dimensional user surface that allows one or two hand input control of computer**

**Patent number:** DE4201934  
**Publication date:** 1993-07-29  
**Inventor:** MAGGIONI CHRISTOPH (DE); WIRTZ BRIGITTE (DE)  
**Applicant:** SIEMENS AG (DE)  
**Classification:**  
- **international:** G06F3/03  
- **european:** G06F3/00B8, G06F3/033A1  
**Application number:** DE19924201934 19920124  
**Priority number(s):** DE19924201934 19920124

**Abstract of DE4201934**

The interactive computer system employs an input facility that responds to the movements of one or both hands of the user to manipulate functions. The process may be based upon either a 2 or 3 dimensional user surface.

The system responds to the gestures provided by hand movement and translates the movements into commands that are then used to control processing. Key features of the motions are extracted and stored.

**ADVANTAGE** - Improves the communication between man and computer.

---

Data supplied from the **esp@cenet** database - Worldwide



①9 BUNDESREPUBLIK  
DEUTSCHLAND



DEUTSCHES  
PATENTAMT

⑫ **Offenlegungsschrift**  
⑩ **DE 42 01 934 A 1**

⑤1 Int. Cl. 5:  
**G 06 F 3/03**

②1 Aktenzeichen: P 42 01 934.6  
②2 Anmeldetag: 24. 1. 92  
④3 Offenlegungstag: 29. 7. 93

DE 42 01 934 A 1

⑦1 Anmelder:  
Siemens AG, 8000 München, DE

⑦2 Erfinder:  
Maggioni, Christoph, Dipl.-Inform., 8000 München,  
DE; Wirtz, Brigitte, Dipl.-Inform., 8000 München, DE

⑤4 **Gestik Computer**

⑤7 Datenverarbeitungssystem, welches durch Ausdrucksformen der menschlichen Gestik gesteuert wird. Es werden Bilder des menschlichen Körpers oder von Körperteilen aufgenommen und aus diesen werden Ausdrucksformen der menschlichen Gestik extrahiert. Diese werden in Befehle zur Steuerung des Datenverarbeitungssystems übersetzt. In Verbindung mit einer graphischen Benutzeroberfläche ist auf diese Weise die Steuerung eines derartigen Datenverarbeitungssystems direkt mit den Bewegungen der menschlichen Hand bzw. beider Hände oder anderer gestischer Ausdrucksformen ohne Zuhilfenahme technischer Eingabemedien, wie z. B. einer Maus, möglich.

DE 42 01 934 A 1

Die Verbesserung der Kommunikation zwischen Mensch und Maschine und die Suche nach neuen Eingabetechniken gewinnt bedingt durch immer komplexer werdende Computeranwendungen immer mehr an Bedeutung. Aus diesem Grunde ist es ein Ziel vieler technischer Entwicklungen, neben den klassischen Eingabegeräten eines Computers, wie z. B. Tastatur, Graphik-Tablett, Maus etc. und in Entwicklung befindlichen Eingabemethoden, wie Sprach- und Handschrifterkennung, neue, der natürlichen räumlichen menschlichen Kommunikation angepaßte Interaktionsmöglichkeiten mit der Maschine zu entwickeln. Hierfür bieten sich die Gestik und Mimik des Menschen auf natürliche Weise an, da diese Ausdrucksformen in der natürlichen Kommunikation zwischen Menschen von großer Bedeutung sind. Um diese Ausdrucksformen des Menschen als Eingabemethode für Computer nutzbar zu machen, müssen Kopf- und Körperbewegungen erkannt und interpretiert werden.

Das bisher einzige am Markt tatsächlich etablierte benutzerfreundliche Eingabemedium für Computer ist die Maus. Betrachtet man den Einfluß dieses Eingabemediums auf die gesamte Entwicklung von Computern bzw. deren Benutzerschnittstellen, so stellt man eine sehr große Steigerung der Benutzerfreundlichkeit und der Effektivität von damit ausgerüsteten Computern fest. Fensteroberflächen, in denen graphische Objekte (Icons) mit der Maus manipuliert werden können, stellen im Vergleich zu den alten, an der Tastatur orientierten Bedienweisen eine völlig neue Art dar, Computer zu benutzen. Dieser qualitative Sprung in der Bedienung ermöglichte eine wesentliche Steigerung der Arbeitsproduktivität, eine deutliche Verringerung der Anlernzeiten und machte viele Computeranwendungen erst möglich.

Die Maus ist ein zweidimensionales Eingabemedium. Im Gegensatz dazu denkt und bewegt sich der Mensch in dreidimensionalen Umgebungen. Seine Gesten und Wahrnehmungen sind räumlich geprägt. Diese Umstände sollten konsequenter Weise in einem dreidimensionalen Eingabemedium und in entsprechenden dreidimensionalen Benutzeroberflächen ihre Entsprechung finden.

Neben der Maus wurde in den letzten Jahren ein weiteres Eingabemedium, welches als Data Glove (Datenhandschuh) bezeichnet wird, bekannt. Diese Data Glove wurde in den Jahren 1984 bis 1987 von der Firma VPL-Research von Thomas G. Zimmermann und L. Young Harvill entwickelt. Der Data Glove übersetzt Bewegungen der Hand und der Finger in elektrische Signale. Zwischen zwei Stofflagen verlaufen fieberoptische Kabel entlang aller Finger und des Daumens. Beide Enden jedes Kabels münden in ein Interface-Board nahe des Handgelenks. Eine lichtemittierende Diode an einem Ende sendet Licht entlang des Kabels zu einem Phototransistor am anderen Ende, der das Licht in ein elektrisches Signal umwandelt. Das Signal gelangt vom Handgelenk zum angeschlossenen Rechner durch ein angeschlossenes Kabel. Je stärker die Finger gekrümmt werden, desto größer ist der Lichtverlust im fieberoptischen Kabel, und desto kleiner damit das entsprechende elektrische Signal. Die Lage und Orientierung der Hand im Raum wird mit Hilfe eines sog. Pollhemus-Sensors (Mc Donnell Douglas Corporation), der auf dem Handrücken montiert wird, bestimmt. Dieser Sensor mißt die Stärke und Orientierung dreier aufeinander senkrecht

stehender künstlich erzeugter Magnetfelder im Raum.

Die Nachteile des Data Glove sind in erster Linie in seinem hohen Preis und in seiner relativ hohen Anfälligkeit gegenüber mechanischen Einflüssen zu sehen. Ferner bestehen erhebliche Akzeptanzprobleme, da der Data Glove nicht immer angenehm zu tragen ist. Ferner schränkt der Data Glove die Beweglichkeit des Anwenders durch Zuführungskabel und sein Gewicht ein. Der Benutzer ist durch den Data Glove allgemein in seiner Aktionsfähigkeit eingeschränkt, da er die Hand nicht für weitere Arbeiten, wie z. B. Schreiben, verwenden kann.

Die Firma Xerox Palo Alto Research Center, der Erfinder der Schreibtischbenutzeroberflächen von Computersystemen, stellte vor kurzem einen Prototyp einer dreidimensionalen Benutzeroberfläche vor. Die Idee ist hierbei, daß sich auf dem Bildschirm Strukturen befinden, welche sich in Echtzeit im dreidimensionalen Raum bewegen, und vom menschlichen visuellen System in ihrer räumlichen Struktur wahrgenommen und interpretiert werden (Mark A. Clarkson, Xerox PARC: "An Easier Interface", Byte Magazine", February 1991, pp. 277-282; George G. Robertson, Stuart K. Card and Jock D. Mackinlay, Xerox PARC: "The Cognitive Coprocessor Architecture for Interactive User Interfaces", Proceedings of the ACM SIGGRAPH Symposium on User Interface Software and Technology, New York 1989, pp. 10-18).

Für eine sinnvolle Arbeit mit derartigen dreidimensionalen Benutzeroberflächen wird ein dreidimensionales Eingabegerät für Computer benötigt. Wünschenswert ist ein dreidimensionales Eingabemedium, welches den Benutzer nicht behindert, so daß er frei arbeiten kann, das vom Benutzer möglichst wenig Anpassung an seinen Rechnerarbeitsplatz fordert, das in der Lage ist, menschliche Metaphern zu erkennen und zu interpretieren und das relativ unabhängig gegenüber Umwelteinflüssen ist.

Der Erfindung liegt die Aufgabe zugrunde, ein Datenverarbeitungssystem anzugeben, welches in der Lage ist, die Ausdrucksformen der menschlichen Gestik direkt in Befehle zu seiner Steuerung umzusetzen. Diese Aufgabe wird durch ein Datenverarbeitungssystem mit Merkmalen nach Anspruch 1 gelöst.

Bei diesem Datenverarbeitungssystem, welches durch Ausdrucksformen der menschlichen Gestik gesteuert wird, sind Mittel vorgesehen, zur Aufnahme von Bildern des menschlichen Körpers oder von Körperteilen, zur Extraktion solcher Ausdrucksformen aus solchen Bildern, und zur Übersetzung solcher Ausdrucksformen in Befehle zur Steuerung des Datenverarbeitungssystems.

Ein derartiges Datenverarbeitungssystem macht die Verwendung technischer Eingabemedien, wie z. B. einer Maus oder einer Tastatur, weitgehend, wenn nicht völlig, entbehrlich. In einer bevorzugten Ausführungsform der Erfindung ist bei dem Datenverarbeitungssystem eine graphische Benutzeroberfläche vorgesehen, deren Bedienungselemente mit Hilfe von Objekten manipuliert werden können, welche extrahierte Ausdrucksformen der menschlichen Gestik darstellen, wodurch das Datenverarbeitungssystem gesteuert wird. Die graphische Benutzeroberfläche kann dabei sowohl zweidimensional als auch dreidimensional sein. Eine Weiterbildung der Erfindung sieht vor, daß extrahierte Ausdrucksformen der menschlichen Gestik durch Objekte auf dieser graphischen Benutzeroberfläche dargestellt werden, welche geeignet sind, Bewegungen und Stellungen der menschlichen Hand oder beider Hände eines Menschen zu visualisieren. Dabei ist es besonders vor-

teilhaft, wenn die graphische Benutzeroberfläche des Datenverarbeitungssystems Strukturmerkmale der realen Arbeitsumgebung eines typischen Anwenders des Datenverarbeitungssystems oder Strukturmerkmale von Gegenständen aus dieser Arbeitsumgebung aufweist, wobei Bedienungselemente dieser Benutzeroberfläche mit Hilfe graphischer Darstellungen von Gegenständen aus dieser Arbeitsumgebung oder Teilen solcher Gegenstände visualisiert werden. Je nach Anwendungsart kann die graphische Benutzeroberfläche dabei z. B. Strukturmerkmale eines Schreibtisches aufweisen. Zusätzlich oder alternativ dazu kann die graphische Benutzeroberfläche auch Strukturmerkmale eines Karteikastens oder mehrerer Karteikasten aufweisen. Für viele Anwendungen ist es vorteilhaft, wenn die graphische Benutzeroberfläche außerdem Strukturmerkmale eines Schriftstücks oder mehrerer Schriftstücke aufweist. Schließlich gibt es eine Reihe von Anwendungen, für welche die graphische Benutzeroberfläche vorteilhaft Strukturmerkmale eines Aktenschanks oder mehrerer Aktenschränke aufweist. Für eine Reihe von Anwendungen ist es vorteilhaft, wenn die Benutzeroberfläche gleichzeitig Strukturmerkmale von mehreren der genannten Arbeitsumgebungen bzw. von mehreren Gegenständen aus einer Arbeitsumgebung aufweist.

In anderen Anwendungen sind graphische Benutzeroberflächen besonders geeignet, welche Strukturmerkmale einer technischen Leitwarte aufweisen. Schließlich ist es in einer Reihe von medizintechnischen Anwendungen vorteilhaft, wenn die graphische Benutzeroberfläche Strukturmerkmale bestimmter medizintechnischer Geräte aufweist.

Die Extraktion von Ausdrucksformen der menschlichen Gestik gelingt besonders zuverlässig und effizient, wenn zu diesem Zweck die Lage von Markierungen auf der Oberfläche des menschlichen Körpers ausgewertet wird. Bei vielen Anwendungen kommt hierfür bevorzugt die Auswertung der Lage von Markierungen auf der Handoberfläche in Betracht. In anderen Anwendungen ist es vorteilhaft, die Blickrichtung eines Anwenders, seine Augenbewegungen, seine Kopfbewegungen, seine Mundbewegungen, seine Mimik oder die Bewegung des Körpers insgesamt auszuwerten. Mit Hilfe eines derartigen Datenverarbeitungssystems ist die intuitive Bedienung eines Computers in verschiedensten Anwendungen möglich. Im Ergebnis wird hierdurch quasi ein neues dreidimensionales Eingabegerät für Computer geschaffen, welches Handgesten und Bewegungen der menschlichen Hand erkennt, verfolgt, interpretiert und dann zur Steuerung von Rechneraktionen verwendet.

Gegenüber bestehenden Eingabemedien wie z. B. einer Maus oder einem Trackball, besitzt die erfindungsgemäße Lösung folgende Vorteile: Sie gewährleistet eine noch intuitivere Bedienbarkeit des Rechners, welche vom Benutzer ein Minimum an Anpassung an seinen Rechnerarbeitsplatz fordert. Die Freiheit des Benutzers wird weitgehend erhalten, da er kein zusätzliches Eingabegerät bedienen muß und deshalb in seiner gewohnten Arbeitsumgebung frei weiterarbeiten kann. Beide Hände des Benutzers können gleichzeitig zur Steuerung des Datenverarbeitungssystems verwendet werden. Die räumliche Welt des Benutzers wird bei Verwendung einer dreidimensionalen Benutzeroberfläche in eine räumliche Interaktionswelt mit dem Rechner abgebildet.

#### Patentansprüche

1. Datenverarbeitungssystem, welches durch Ausdrucksformen der menschlichen Gestik gesteuert wird, mit Mitteln

- zur Aufnahme von Bildern des menschlichen Körpers oder Teilen davon,
- zur Extraktion solcher Ausdrucksformen aus solchen Bildern, und
- zur Übersetzung solcher Ausdrucksformen in Befehle zur Steuerung des Datenverarbeitungssystems.

2. Datenverarbeitungssystem nach Anspruch 1 mit einer graphischen Benutzeroberfläche, auf der erste Objekte sichtbar sind, welche Bedienungselemente der Benutzeroberfläche darstellen, und auf der zweite Objekte sichtbar sind, welche extrahierte Ausdrucksformen der menschlichen Gestik darstellen, wobei erste Objekte mit Hilfe von zweiten Objekten manipuliert werden können, wodurch das Datenverarbeitungssystem gesteuert wird.

3. Datenverarbeitungssystem nach Anspruch 2 mit einer zweidimensionalen graphischen Benutzeroberfläche.

4. Datenverarbeitungssystem nach Anspruch 2 mit einer dreidimensionalen graphischen Benutzeroberfläche.

5. Datenverarbeitungssystem nach einem der Ansprüche 2 bis 4, bei dem extrahierte Ausdrucksformen der menschlichen Gestik durch Objekte dargestellt werden, welche geeignet sind, Bewegungen und Stellungen der menschlichen Hand oder beider Hände eines Menschen zu visualisieren.

6. Datenverarbeitungssystem nach einem der Ansprüche 2 bis 5, dessen graphische Benutzeroberfläche Strukturmerkmale der realen Arbeitsumgebung eines typischen Anwenders des Datenverarbeitungssystems oder Strukturmerkmale von Gegenständen aus dieser Arbeitsumgebung aufweist, wobei Bedienungselemente dieser Benutzeroberfläche mit Hilfe graphischer Darstellungen von Gegenständen aus dieser Arbeitsumgebung oder Teilen solcher Gegenstände visualisiert werden.

7. Datenverarbeitungssystem nach einem der Ansprüche 2 bis 6, dessen graphische Benutzeroberfläche Strukturmerkmale eines Schreibtisches aufweist.

8. Datenverarbeitungssystem nach einem der Ansprüche 2 bis 7, dessen graphische Benutzeroberfläche Strukturmerkmale eines Karteikastens oder mehrerer Karteikasten aufweist.

9. Datenverarbeitungssystem nach einem der Ansprüche 2 bis 8, dessen graphische Benutzeroberfläche Strukturmerkmale eines Schriftstücks oder mehrerer Schriftstücke aufweist.

10. Datenverarbeitungssystem nach einem der Ansprüche 2 bis 9, dessen graphische Benutzeroberfläche Strukturmerkmale eines Aktenschanks oder mehrerer Aktenschränke aufweist.

11. Datenverarbeitungssystem nach einem der Ansprüche 2 bis 10, dessen graphische Benutzeroberfläche Strukturmerkmale einer technischen Leitwarte aufweist.

12. Datenverarbeitungssystem nach einem der Ansprüche 1 bis 11, dessen graphische Benutzeroberfläche Strukturmerkmale medizintechnischer Geräte aufweist.

13. Datenverarbeitungssystem nach einem der vor-

hergehenden Ansprüche, welches die Lage von Markierungen auf der Oberfläche des menschlichen Körpers zur Extraktion von Ausdrucksformen der menschlichen Gestik auswertet.

14. Datenverarbeitungssystem nach einem der vorhergehenden Ansprüche, welches die Lage von Markierungen auf der Handoberfläche zur Extraktion von Ausdrucksformen der menschlichen Gestik auswertet.

15. Datenverarbeitungssystem nach einem der vorhergehenden Ansprüche, welches die Blickrichtung eines Anwenders auswertet.

16. Datenverarbeitungssystem nach einem der vorhergehenden Ansprüche, welches die Augenbewegungen eines Anwenders auswertet.

17. Datenverarbeitungssystem nach einem der vorhergehenden Ansprüche, welches die Kopfbewegung eines Anwenders auswertet.

18. Datenverarbeitungssystem nach einem der vorhergehenden Ansprüche, welches die Mundbewegungen eines Anwenders auswertet.

19. Datenverarbeitungssystem nach einem der vorhergehenden Ansprüche, welches die Mimik eines Anwenders auswertet.

20. Datenverarbeitungssystem nach einem der vorhergehenden Ansprüche, welches die Körperbewegung eines Anwenders auswertet.

30

35

40

45

50

55

60

65